

聴覚過敏を有する自閉スペクトラム症者の音声に対する主観的経験を予測するモデルの支援への応用に向けた分析

Considering for Application to Support of a Machine Learning Model Which Predicts Subjective Auditory Experience of Autism Spectrum Disorders with Auditory Hypersensitivity

市川 樹[†] 長井 志江[‡] 國吉 康夫[‡] 和田 真[†]
Itsuki Ichikawa Yukie Nagai Yasuo Kuniyoshi Makoto Wada

1. はじめに

発達障害の一種である自閉スペクトラム症(ASD)当事者の多くが示す生活上の問題として、特定の音に対し強いストレスを感じる、大きな音や多くの音がある環境で疲労するといった聴覚過敏の症状がある。ASD 当事者の多くは視覚、触覚といった他の感覚モダリティにも問題を抱えているが、それらと比較しても聴覚に関する困りごとが特に強いことが報告されている[1]。現在でも当事者は耳栓やイヤーマフといった器具を用いる、周囲への配慮を求めるといった形で対処することができるが、聴覚過敏性の問題がどのように表出するかは当事者個人々の特性/生活環境によって異なり、また同一個人でも体調などの要因で変化しうるため一括的な対策だけでは対応しきれない場合があると考えられる。

発表者はこの問題に対処すべく、これまで ASD 当事者が音をどう感じているのか、即ち入力される音声刺激に対する主観的体験が本来の音声刺激からどう変化しているのかに関するデータを収集し、外界から来る音声刺激に対して主観的体験の変化を予測させることを目的としてディープニューラルネットワーク(DNN)を用いた機械学習モデルを開発してきた[2]。本稿では、将来的なモデルの当事者支援への将来的な応用に向けた実験およびその分析結果について報告する。第 2 章では先行研究について、第 3 章で主観的聴覚体験予測モデルの開発のために行われたデータ収集実験について、第 4 章で実際のモデルの構造と学習条件及び結果(各タスクに対する精度)、入力データの聴覚体験予測結果に対する寄与率、および入出力データ間の関係性に関するクラスタリングと実験参加者群間・個人間での比較の可能性について解説し、第 5 章でまとめと今後の展望を述べる。

2. 先行研究

本研究の主観的経験予測モデルの開発手法は長井らによる ASD 視覚体験シミュレータに関する研究[3]のものを参考にした。この研究では ASD 診断を受けた成人 22 名に対して様々な場面を表す動画刺激を提示し、6 種類の視覚フィルタ(砂嵐状ノイズ、コントラスト強調など)によって動画を変化させ、自らが過去に同様の場面でどのような見え方をしていたかを再現させるタスクを実施した。各種フィルタの設定量が即ち ASD 当事者の主観的視覚体験が本来の刺激に対してどれだけ変化しているかを表すものとして動画特徴量との正準相関分析を行った結果、「動画の輝度とコントラスト強調・高輝度化フィルタの設定値が相関する」といった一定の傾向があることが示唆され、ASD 当事者の知覚変化の過程をモデル化することで非 ASD 者が ASD 当

事者の視覚体験を体感することができるシミュレータを開発することに成功した。同様の感覚体験変化のモデル化の試みは聴覚モダリティに対しても行われている[4]。

発表者はこれらの研究をもとに、DNN を用いた感覚刺激と主観的経験の変化の結びつけを試みてきた。DNN は非線形的な関係を取り扱うことができるため回帰分析では対応しきれなかった関係性に対応できることが期待され、また刺激データからモデル自身が自動的に特徴量を得るため ASD 当事者の感覚体験の変化に関わる未知の要素の発見にもつながりうる。

3. データ収集実験

主観的経験予測モデルの開発のため、(1)実験参加者に対して音声刺激を提示し、(2)参加者がそれに対して各種の聴覚フィルタをかけて音声を変化させる実験タスクを実施した。タスクは音声を変化させるべき方針の違いで再現/緩和の 2 種類が存在し、実験参加者はそれら両方に参加した。実験参加者は 2023 年 5 月時点で ASD 者(ASD の診断を受けたグループ)が 14 名、定型発達者が 13 名で合計 27 名であった。実験プログラムは iPad(第 9 世代)にインストールされており、実験参加者は防音室内で着席し Bluetooth ヘッドホン(Logicool ZoneVibe 100)を装着して実験を受けた。実験開始前に、参加者は年齢、性別、身体的な疲れ具合(1-7 の値で回答)、精神的な疲れ具合(1-7 の値で回答)、睡眠時間(時間・分を指定し回答)について回答した。

3.1 再現タスク

このタスクでは、実験参加者は音声刺激に対し「刺激と同様の場面で自分がどのような聞こえ方をしていたかを再現する」ために聴覚フィルタを用いることを求められる。こ

表 1: 再現タスクのフィルタ

名称	効果
音の増幅	音を全体的に大きくする。
低音の増幅	低音(20-200Hz)を大きくする。
中音の増幅	中程の音(200-2000Hz)を大きくする。
高音の増幅	高音(2000-20000Hz)を大きくする。
ノイズ	ホワイトノイズを添加する。
耳鳴り	耳鳴り様のサイン音を添加する。強さ、高さを設定可能。
おさえこみ	一定範囲の高さの音の音量を抑える。範囲の中心、幅を設定可能。
エコー	音が反響するような感覚を付与する。
うねり	音がうねるような感覚を付与する。強さ、速さを設定可能。
水中音	音をくぐもらせ、水中で聴いている様な感覚を付与する。

表 2: 緩和タスクにおける、再現タスクからのフィルタの変更点

変更されるフィルタ	
名称	効果
音の増幅 →全体変更	音を全体的に増幅させる以外に、小さくすることも可能になる。
低音の増幅 →低音変更	周波数が低い音(20-200Hz)を大きく、あるいは小さくする。
中音の増幅 →中音変更	周波数が中程の音(200-2000Hz)を大きく、あるいは小さくする。
高音の増幅 →高音変更	周波数が高い音(2000-20000Hz)を大きく、あるいは小さくする。
追加されるフィルタ	
ピッチシフト	音の高さを低くあるいは高くする。
変化抑制	音が急激に大きくなった際にフィルタが作動し、音量を抑制する。設定項目は(1)音がどの程度急激に大きくなったらフィルタを作動させるか、(2)フィルタ作動時に音をどの程度小さくするか、(3)音が一旦小さくなった後、どの程度時間をかけて元の音量に戻すか、の3つである。



図 1: 実験プログラムの画面

(左: フィルタ設定画面、右: アンケート画面)

のタスクで実験参加者が用いる聴覚フィルタの一覧を表 1 にて示す。フィルタの種類に関しては先行研究[4]で行われていた実験タスクのものに加え、ASD 当事者が主に高音に対して苦手意識を感じる人が多い[1]ことを踏まえ周波数帯ごとに音量を増幅できるものを加えた。

フィルタ設定が完了したあとは、参加者は音声に対してどう感じたかを回答する。の設問は(1)音が全体としてどの程度辛かったか、(2)音に対して痛みを感じるような感覚がどの程度あったか、(3)音によって気が散る・注意力がなくなる感覚がどの程度あったか、(4)音によってどの程度不安感を感じたか、(5)音によって聞きたい音が聞こえなくなる感覚がどの程度あったか、(6)過去にこの音声と同じような場面を具体的にどこで経験したか、(7)音に対して感じたことに対する自由記述、となっている。(1)から(5)についてはどの程度当てはまるかを 1-7 で評価し、(6)については「実際に遭遇したことがある」「TV や動画 (フィクション含む) で見た事がある」「これ自体は経験がないが似た

表 3: 実験で使用した音声データの一覧表

種類	内容
学習用 (30 件)	救急車サイレン、電子レンジの作動音、マウスクリック、冷蔵庫作動音、海辺の波の音、エアコンの音、学校のチャイム、教会の鐘、コップに水を注ぐ音、パトカーサイレン、カラスの鳴き声、シャワー、シンクに水が垂れる音、せき払い、セミの鳴き声、滝、電動ノコギリ、電話呼び出し音(電子音)、目覚まし時計(ベル)、目覚まし時計(電子音)、赤ん坊の鳴き声、電車内からの走行音、揚げ物の油の音、スポーツ試合中、バスの通過音、犬の吠え声、バイク走行音、車のエンジン音、カモメの鳴き声、音のついた横断歩道
学習用/予備 (10 件)	人の多い通り、ドライバー、エレベーター音声、街頭アナウンス、ドラムロール、静かな部屋での動画の読み上げ、咀嚼音、水鳥の鳴き声、いびき、雷雨
テスト用 (12 件)	プリンタ作動音、人の少ない通り、親子連れがいる屋外、ドアの開け閉め、ビニール袋にものを詰める音、足音、花火大会、スーパーの店内、咳き込み、食器の出し入れ、群衆が怒る声、蛍光灯の音
テスト用/予備 (10 件)	工事現場、葉っぱの上を歩く音、スタジアムの群衆の声、レストラン内の声、川の流れの音、銃声、スーパーマーケット店内(2)、風船を捻る音、キーボードの音、パチンコ店の側の音

ような場面を思い出した」「全く経験がない」から選択する。

3.2 緩和タスク

再現タスクを実施した後、続けて同一の音声刺激に対して緩和タスクを行う。このタスクでは、参加者は「刺激がこのように聞こえていれば、ずっと聞こえていてもストレスを感じることがなくなるように変化させる」ために聴覚フィルタを用いることを要求される。ここでも再現タスクと同様のフィルタを使用するが、一部については機能が変更されており、また新たなフィルタも追加されている。変更・追加されたフィルタについては表 2 に示す。変更されたフィルタを用いることで音を完全に無音にすることも可能だが、実験参加者には無音にはせず、元の音がなんであるか判別できる範囲で変化させるよう指示した。なお、緩和タスクで流れる音声刺激には再現タスクにおけるフィルタの設定量が反映されており、本来の音からの主観的経験のずれを踏まえた形でタスクを行うことになる。

フィルタ設定終了後、参加者は変化させた音声に対してどう感じたかを回答する。設問は「音声刺激と同様の場面をどこで経験したか」がないことを除き、再現タスクのもの

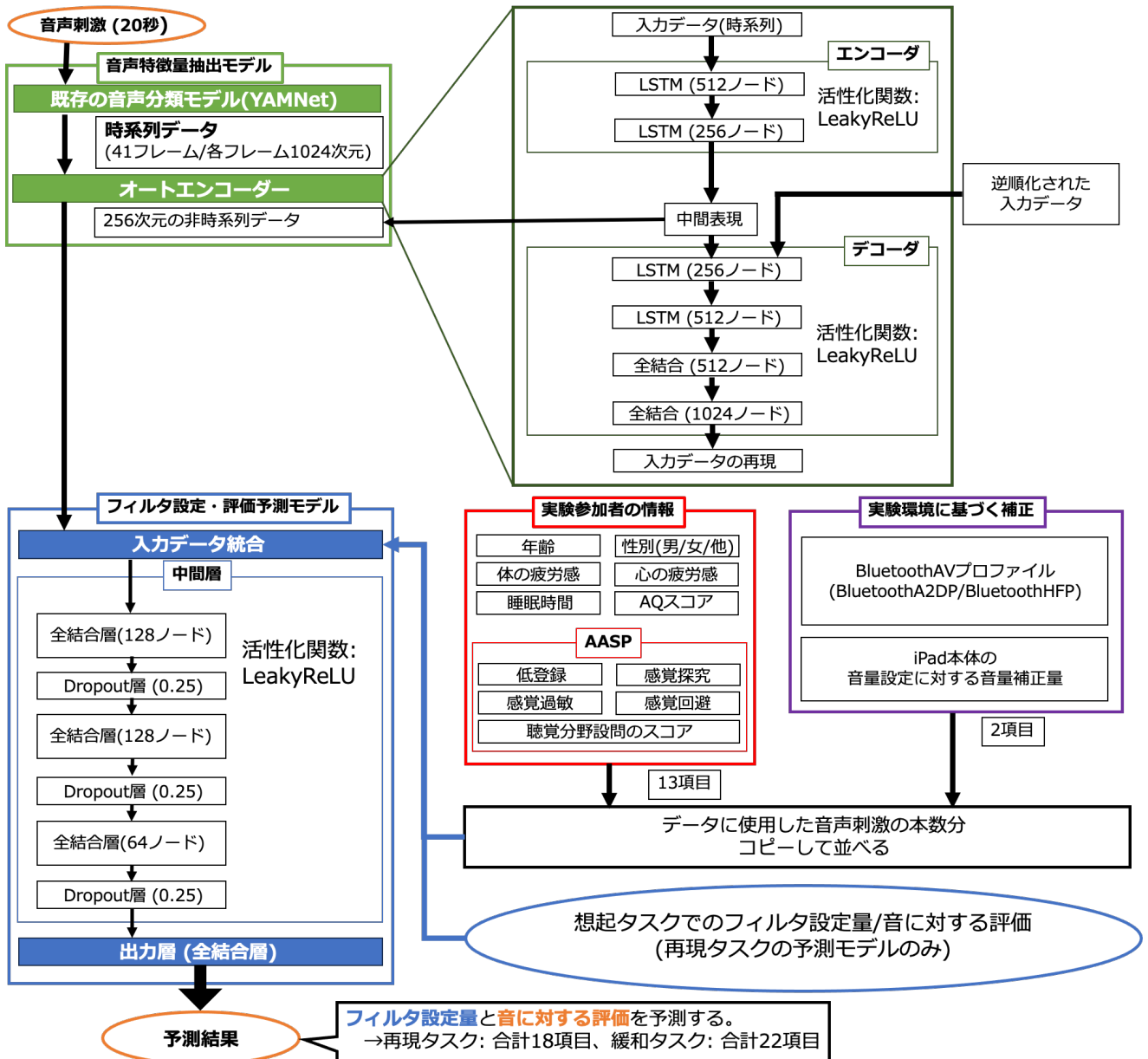


図 2: 本学習で用いた機械学習モデルの模式図

のと同様である。回答後は次の音声刺激に対して再び再現・緩和タスクを実施した。

3.3 使用した音声刺激

実験参加者は合計で45件の音声刺激(各20秒)に対し、先述した再現・緩和の2種類のタスクを行った。これらのうち30件に対するデータは深層学習モデルの学習のため、12件はモデルの精度テストのために用いられる。残る3件については実験参加者が正しく自分の感覚を再現できているかどうかを評価するために、学習用の音声と同一のものについて再度回答させた。これらのデータは将来的に学習内容の選別に用いられる予定だが、本発表においては再現精度を問わず全参加者のデータをモデルの学習・検証対象とした。

また、参加者は音声刺激に対して過去に類似の経験をしている必要があったため、そのような経験のない音刺激が提示された場合、それをスキップした上で予備の音刺激を提示することも可能とした。表3に予備も含めて使用された音声刺激の一覧を示す。なお、提示順は実験参加者ごとにシャッフルされている。

3.4 実験プログラム

本実験に用いたプログラムの画面を図1に示す。実験参加者はフィルタ設定画面中央下部のフィルター一覧(左画面中央下、アイコンが並んでいる部分)を指でスライドさせることで設定中のフィルタを変更することができ、フィルター一覧の下にあるスライダーを動かすことで設定中フィルタのパラメータを変化させて、リアルタイムに音声を変化

表 4: モデルの学習条件
(再現タスク・緩和タスクで共通)

項目	説明・数値
最適化関数	Adam(学習率=0.001, $\beta_1=0.9, \beta_2=0.999$)
損失関数	平均二乗誤差
訓練用データ中の検証用データ割合	0.15
エポック数	200
バッチサイズ	32

表 5: 再現タスクにおける精度, および各データセット中でフィルタの設定が変更された件数

フィルタ設定量(精度は Pearson 相関係数)			
名称	学習用/設定変更件数(全体 809 件)	テスト用/設定変更件数(全体 324 件)	精度
音の増幅	532	207	0.508(***)
低音の増幅	453	162	0.390(***)
中音の増幅	355	166	0.503(***)
高音の増幅	519	183	0.498(***)
ノイズ	167	67	0.372(***)
耳鳴り/強さ	83	35	0.264(***)
耳鳴り/高さ	135	25	0.226(***)
おさえこみ/幅	266	83	0.174(***)
おさえこみ/中心	162	35	0.0858
エコー	221	78	0.372(***)
うねり/強さ	117	28	0.337(***)
うねり/速さ	96	18	0.552(***)
水中音	113	18	0.163(***)
音に対する評価(精度は Spearman 相関係数)			
全体的なつらさ			0.563(***)
痛みの感覚			0.475(***)
気が散る感覚			0.563(***)
不安になる感覚			0.495(***)
聞きたい音が聞けない			0.540(***)

(***): $p < 0.001$, (**): $p < 0.01$, (*): $p < 0.05$

させることができる。設定中、音声は常にループしている。設定が完了すると図 1 右側に示されたアンケート画面に移行し、回答を終えると再現タスクであれば同じ音声に対する緩和タスクへ、緩和タスクであれば次の音声の再現タスクに移行する。

4. フィルタ設定結果・評価予測モデルとその学習結果

4.1 モデル解説

本研究で用いた機械学習モデルの模式図を図 2 に示す。本研究で使用したものは音声特徴量を抽出するモデルとフィルタ設定結果・音に対する評価を予測するモデルの 2 つに分かれている。音声特徴量抽出モデルはまず既存の音声

表 6: 緩和タスクにおける精度, および各データセット中でフィルタの設定が変更された件数

フィルタ設定量(精度は Pearson 相関で計算)			
名称	学習用/設定変更件数(全体 809 件)	テスト用/設定変更件数(全体 324 件)	精度
全体変更	465	176	0.519(***)
低音変更	497	187	0.211(***)
中音変更	436	165	0.277(***)
高音変更	522	189	0.286(***)
ピッチシフト	297	79	0.447(***)
変化抑制/作動しきい値	246	89	0.620(***)
変化抑制/作動時抑制量	261	88	0.511(***)
変化抑制/作動後の戻りの長さ	87	39	0.387(***)
ノイズ	80	10	0.0850
耳鳴り/強さ	30	4	0.0224
耳鳴り/高さ	96	17	0.694(***)
おさえこみ/幅	312	113	0.587(***)
おさえこみ/中心	209	61	0.315(***)
エコー	128	38	0.0297
うねり/強さ	53	10	0.119(*)
うねり/速さ	40	4	0.0297
水中音	116	35	0.225(***)
音に対する評価(精度は Spearman 相関で計算)			
全体的なつらさ			0.596(***)
痛みの感覚			0.498(***)
気が散る感覚			0.633(***)
不安になる感覚			0.586(***)
聞きたい音が聞けない			0.597(***)

(***): $p < 0.001$, (**): $p < 0.01$, (*): $p < 0.05$

認識モデルである YAMNet[5]を用いて音声データから特徴量を抽出し、そこからさらにオートエンコーダー(AE)[6]を用いて 256 次元の非時系列データに削減するものである。ここで AE とは、入力と出力に同じデータを指定して学習することによってデータの次元を削減した特徴量を得るために用いられる DNN の一種で、学習が完了した後に次元削減を行うまでの部分: エンコーダ部のみをモデルから抜き出し、そこからの出力結果を用いることで任意のデータの特徴量抽出に用いることが可能である。ここでは時系列データに対する AE として、元データを逆順化したものをデコーダ部: 削減したデータから元のデータを再現する部分に元データの次元削減された特徴量と合わせて入力するタイプのモデルを用いている[7]。主観的体験予測モデルは (1)抽出された音声特徴量, (2)実験参加者情報(年齢, 性別, 身体・精神の疲労感, 睡眠時間, ASD 特性の強さを表す自閉症スペクトラム指数(AQ)[8]のスコア, 感覚刺激への反応傾向を表す青年・成人感覚プロフィール(日本版 AASP)[9]のスコア: 低登録・感覚探究・感覚過敏・感覚回避の総合得点4種類に加え聴覚分野の設問11問の合計点数を使用), (3)実験環境に基づく補正量(BluetoothAV プロファイル

表 7: 想起タスクの主観的体験予測モデルにおける音声特徴量(寄与率±0.050 以上)と AQ/AASP スコアの音に対する評価への寄与率

名称(音声特徴量 に関しては 256 次元中の番号)	寄与率(DeepLIFT)				
	全体的なつらさ	痛みの感覚	気が散る感覚	不安になる感覚	聞きたい音が聞けない
音声特徴量_64	-0.162	-0.109	-0.132	-0.116	-0.128
音声特徴量_79	0.123	0.076	0.108	0.091	0.098
音声特徴量_97	-0.131	-0.094	-0.131	-0.082	-0.126
音声特徴量_154	-0.262	-0.182	-0.272	-0.201	-0.227
音声特徴量_171	0.106	0.065	0.136	0.075	0.104
AQ スコア	0.087	0.074	0.095	0.004	0.083
AASP/低登録	0.030	0.028	0.020	0.017	0.020
AASP/感覚探究	0.014	0.028	0.045	0.000	0.081
AASP/感覚過敏	-0.085	-0.062	-0.124	-0.057	-0.121
AASP/感覚回避	0.136	0.119	0.174	0.101	0.190
AASP/聴覚部分のスコア	0.086	0.068	-0.012	0.085	0.025

が“BluetoothA2DP”か“BluetoothHFP”か、iPad 本体の音量設定に対しなるべく実際の音量が同等に聞こえるようにするための音量補正量)の3つに基づきフィルタ設定量と音に対して感じた印象(再現タスクのモデルの場合のみ)を予測するものである。なお学習に際し、入力データのうちの実験参加者情報、出力となるフィルタ設定量・音に対する評価は 0-1 の範囲に正規化している。これらのモデルは Python 向け機械学習ライブラリ TensorFlow[10](既存音声認識モデルによる特徴量抽出、オートエンコーダーの部分)および PyTorch[11](主観的体験予測モデルの部分)を用いて開発された。モデルの学習データは第 3 章で収集したデータを用いた。学習データは 27 人×30 件で 810 件となるが、1 件データの欠損があったため 809 件となった。詳細な学習条件を表 4 として示す。

4.2 精度

この項では再現・緩和タスクそれぞれに対して学習したモデルの予測精度について示す。予測精度は各参加者につきそれぞれ 12 件回答した精度テスト用データ(全体では 27 名×12 件で 324 件となる)について学習済みのモデルに予測させ、その値と真の値との間で相関係数を計算したものとなる。この時フィルタ設定量に関しては Pearson 相関を用い、一方で音に対する評価については順序尺度となるため Spearman 相関を用いている。

4.2.1 再現タスク – 主観的体験予測モデル

再現タスクにおけるモデルは、音に対して生じる主観的な体験を予測するモデルとなる。

精度、およびフィルタについてタスク中で設定値が初期値から変更されたデータの件数を表 5 に示す(設定値が初期値である場合、そのフィルタが表す主観的経験の変化は生じていないこととなる)。音量の増幅に関わるフィルタである音の増幅、低音の増幅、中音の増幅、高音の増幅についてはデータセットにおいて概ね全体の半数以上で設定されており、主観的に音が強調される傾向があることがうかがえる。一方、耳鳴り、うねり、水中音に関しては使用されたケースが少なかった。

4.2.2 緩和タスク – 緩和内容予測モデル

緩和タスクにおけるモデルは、音に対するストレスを感じさせないように緩和するためにどのようなフィルタをかければ良いか、またその結果音に対する評価がどのように緩和されるかを予測するモデルとなる。

緩和タスクにおける精度と各フィルタの設定値が初期値から変更された件数を表 6 として示す。こちらも音量の変更に関するフィルタは設定が変更されている割合が大きく、次いでおさえこみ、変化抑制、ピッチシフトがよく使用されていた。精度に関しては、全体変更は再現タスクに近い値であったが、それ以外の音量変更に関わる 3 つについては低下していた。

4.3 再現タスクにおける音声特徴量およびコンディションの音への評価への寄与率

ASD 当事者がどのような要素を持つ音に対してストレスを感じる、あるいは心身の疲れ・睡眠時間といったコンディションの問題が聴覚過敏性にどのように影響するかを解明することで、具体的な支援の方針の考案につなげていくことができると考えられる。本研究では学習が完了した再現タスクのフィルタ設定結果・評価予測モデルに対し、AI の出力に対する根拠を人間が解釈可能な形で提示する技術である説明可能 AI(XAI)の手法を用いることで音刺激・コンディションと音へのストレスの関係を求める。PyTorch 上で作成された機械学習モデルに対し、Captum[12]を用いることで XAI の様々な手法を適用し出力値に対する各入力変数の寄与率を計算することが可能である。今回は Captum に実装されている手法のうち DeepLIFT[13]を用い、想起タスクの主観的体験予測モデルについて出力(音に対する評価 5 種)に対する入力変数の寄与率を計算した。サンプルとして訓練用データセットからランダムに 64 件のデータを抽出して用いている。

計算結果として、256 次元の音声特徴量のうち特に寄与率が高かったもの(各項目±0.050 以上)5 種類と AQ/AASP スコアの音に対する評価における寄与率を表 7 に、それら 5 種類の音声特徴量の使用した各動画における値の分布をグ

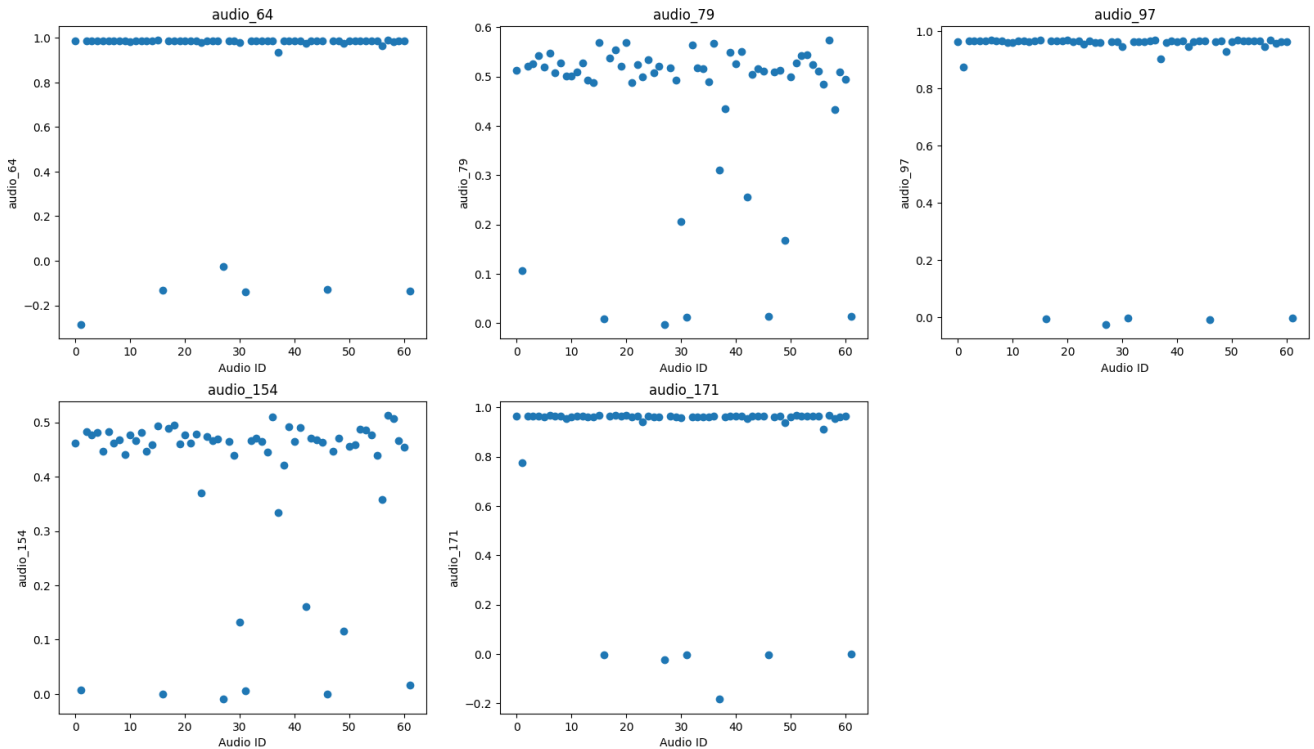


図 3: 表 7 において用いた音声特徴量の、実験で使用した各動画における値の分布

ラフ化したものを図 3 として示す。音声特徴量についてはオートエンコーダーによって抽出される特徴量が疎となったため分布が偏っており、音声刺激ごとの差別化が難しいため、他の寄与率がより低い特徴量についても傾向が似ているものをまとめて勘案するかモデルの構造を再考する必要があるものと考えられる。

一方で AQ/AASP スコアについては AASP/感覚過敏がどの項目にも負の寄与をして AASP/感覚回避は正の寄与をしていたこと(即ち、感覚過敏/感覚回避傾向が強いほど音に対する辛さを表す各項目が低くなる/高くなる)が注目される。AASP において感覚過敏は感覚に対する閾値が低さ、即ち受動的な指標であり、一方で感覚回避は刺激を避けようとする傾向、即ち能動的な行動を表す指標である。そのため、特定の音刺激に対して能動的に注意が向くことが音に対する辛さに結びつく可能性がある。また、ASD 特性の強さを示す AQ スコアは「不安になる感覚」を除くすべての項目で正の寄与を示している。先行研究では聴覚過敏性の強さと相関していた聴覚パラメータ(耳小骨筋反射)と同じく ASD 特性を表す別指標(SRS-2 対人応答性尺度)と相関していたとするものもあり[14]、ASD 特性の強さと聴覚過敏性の強さの関連性を示す結果であると考えられる。

4.4 音声特徴量-主観的体験の関係のクラスタリング

ASD 当事者が抱える聴覚過敏性の問題は個々人で異なるが、具体的にどのような違いがあるかについて扱った研究は現状ではまだ不足している。再現タスクに対する主観的経験予測モデルについて、入出力間の関係性は音声刺激に対してどのような感じ方をするか(フィルタ設定量、音に対する評価設定)を示すものであると考えられ、それらを

クラスタリングすることで聴覚過敏性の傾向を発見することが期待される。

先行研究において、分類タスクで出力層からひとつ手前の中間層のパラメータを抽出してクラスタリングすることによりデータのカテゴリ別に固まった位置に配置されることが確かめられており[15]、回帰タスクを扱う本研究においても同様のパラメータ抽出・クラスタリングを行うことで入出力間の関係性をマッピングできるものと考えられる。

本研究では多次元のデータを 2 次元に落とし込んでクラスタリングするための手法として t-Distributed Stochastic Neighbor Embedding(t-SNE)[16]を用いた。t-SNE は高次元における各データ間の関係性を極力保つように次元圧縮を行うことが可能な手法である。クラスタリングは図 2 における出力層前最後の中間層(64 ノード)内のパラメータに対して行った。t-SNE のハイパーパラメータである perplexity(あるデータについて、近傍のデータのうちの程度を有効なもののみならず)は 10、処理の反復回数は 5000 回とした。

クラスタリングを実施した結果について図 4 に示す。(a) では ASD 群/対称群の差異に注目し AQ の数値によるプロットへの色づけを行ったところ、概ね左→右、あるいは中心から上下に向かうグラデーションが生じていることが確認できる。加えて、ASD 群であっても AQ スコアが低い場合対称群のデータが多く含まれる位置に含まれ、逆に対称群でも AQ スコアが高めの参加者のデータは ASD 群が主に占める位置にあった。これらは、AQ が ASD 診断の有無を問わず連続した値であることを踏まえれば尤もらしい結果であると考えられる。また図 4(b)~(d)では異なる実験参加者 3 名にそれぞれについてプロットをハイライトした結果、1 人ごとに明確に離れた位置にプロットが分布しているこ

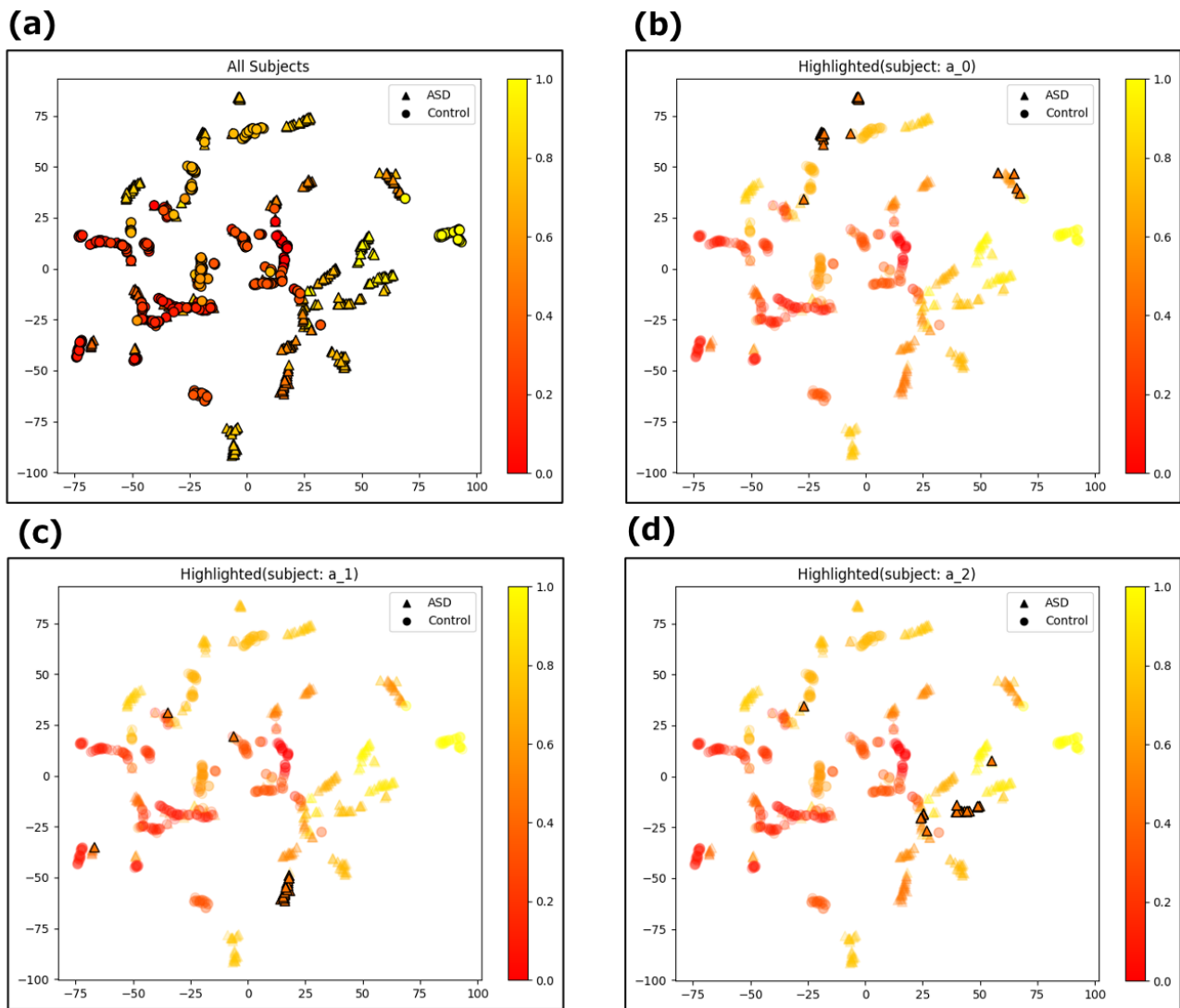


図 4: 音声特徴量-主観的体験の関係のクラスタリング結果 (a)全体 (b)-(d)ASD 群実験参加者 1 名ずつのデータのハイライト (プロットの色は AQ スコアの高低を示す)

とが確かめられた。このことから、個人間での特性の比較も可能であると考えられる。

5. おわりに

本稿では発表者がこれまで開発を進めてきた ASD 当事者の音に対する聴覚過敏性を予測する機械学習モデルについて、入力音や個人特性の主観的体験に対する寄与率の計算・関係性のクラスタリング、また音を自ら緩和できるとしたらどのように緩和するかに関する実験タスクを試みた。これらの試みは聴覚過敏性の問題を抱える ASD 当事者がストレスなく過ごせる環境を作るための方針の提案や、モデルを携帯可能な機器に移植した上で過敏性の問題を生じうる環境におかれた際の対処法のサジェストを行うアプリケーションを開発することにつながるものである。また、緩和タスクにおけるフィルタ設定・音に対する評価の予測モデルを使用して、外から入ってくる音刺激に対しモデルが予測したフィルタを適用することによって自動的に音に対するストレスの緩和を試みるシステムを開発することも現在検討中である。

今後は機械学習に関する既存の知見を踏まえたモデルの再検討や必要に応じた追加のデータ収集を通じてこれらの目的に足る精度を発揮するモデルを開発してゆき、最終的には実際の当事者の QoL 向上のための支援に繋げていくことを目標とする。

謝辞

本研究の実施にあたり発表者の所属研究室の技術協力員である石井亨視さんにご協力いただきました。研究全体の進行に関して、東京大学国吉・新山研究室スタッフの方々からご助言いただきました。本研究は科研費(21H05053)の支援により実施されました。

参考文献

- [1] Wada, M., Hayashi, K., Seino, K., Ishii, N., Nawa, T., & Nishimaki, K. (2023). Qualitative and quantitative analysis of self-reported sensory issues in individuals with neurodevelopmental disorders. *Frontiers in Psychiatry*, 14, 109.
- [2] 市川樹, 長井志江, 國吉康夫, 和田真. (2022). “自閉スペクトラム症(ASD)者の聴覚過敏性予測モデルへの既存モデル組み込みの検討”, 日本発達神経科学会第 11 回学術集会.

- [3] 長井志江, 秦世博, 熊谷晋一郎, 綾屋紗月, & 浅田稔. (2015). 自閉スペクトラム症の特異な視覚とその発生過程の計算論的解明: 知覚体験シミュレータへの応用. *日本認知科学会第32回大会発表論文集*, 32-40.
- [4] Hsieh, J. J., Nagai, Y., Kumagaya, S. I., Ayaya, S., & Asada, M. (2022). Atypical Auditory Perception Caused by Environmental Stimuli in Autism Spectrum Disorder: A Systematic Approach to the Evaluation of Self-Reports. *Frontiers in Psychiatry*, 13.
- [5] Tensorflow (2020). YAMNet, <https://github.com/tensorflow/models/tree/master/research/audioset/yamnet> (Accessed: June 5, 2023)
- [6] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- [7] Srivastava, N., Mansimov, E., & Salakhudinov, R. (2015, June). Unsupervised learning of video representations using lstms. In *International conference on machine learning* (pp. 843-852). PMLR.
- [8] 若林明雄, & 東條吉邦. (2004). 自閉症スペクトラム指数 (AQ) 日本語版の標準化 高機能臨床群と健常成人による検討. *心理学研究*, 75(1), 78-84.
- [9] 萩原拓, 岩永竜一郎, 伊藤大幸, & 谷伊織. (2015). AASP 青年・成人感覚プロファイルユーザーマニュアル 日本文化科学社.
- [10] <https://www.tensorflow.org/> (Accessed: June 6, 2023)
- [11] <https://pytorch.org/> (Accessed: June 6, 2023)
- [12] <https://captum.ai/> (Accessed: June 6, 2023)
- [13] Shrikumar, A., Greenside, P., & Kundaje, A. (2017, July). Learning important features through propagating activation differences. In *International conference on machine learning* (pp. 3145-3153). PMLR.
- [14] Ohmura, Y., Ichikawa, I., Kumagaya, S., & Kuniyoshi, Y. (2019). Stapedial reflex threshold predicts individual loudness tolerance for people with autistic spectrum disorders. *Experimental brain research*, 237, 91-100.
- [15] Susan, S., & Malhotra, J. (2020, February). Learning interpretable hidden state structures for handwritten numeral recognition. In *2020 4th International Conference on Computational Intelligence and Networks (CINE)* (pp. 1-6). IEEE.
- [16] Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(11)

† 国立障害者リハビリテーションセンター研究所
National Rehabilitation Center for Persons with Disabilities
‡ 東京大学 The University of Tokyo